

Parameterized database drill-through

Field of the Invention

The present invention relates to interactively searching a database in such a manner that it is quick and easy to search, drill down, drill-up and drill across a data collection, presenting the user with summary information using multiple independent hierarchical category taxonomies of the data collection. The present invention also relates to retrieving information from a database based on content aggregation, management and distribution.

Background of the Invention

The field of business applications of computer technology has seen many important changes over the last few years. With steadily growing computational power and data storage capacities of computer systems used for business data processing, the interest of the business community has shifted from transactional data management systems known as on-line transaction processing (OLTP) systems, mostly supporting day-to-day business operations, and from relatively simple business data processing systems, towards sophisticated business management systems, such as online analytical processing (OLAP) and business intelligence (BI) systems, integrating at the enterprise level many facets and functions of the business, including planning, manufacturing, sales and marketing. Collectively, these are often termed decision support systems.

Among various alternative approaches to business data management and analysis developed over the last few years, many are related to data warehousing. A data warehouse can be defined broadly as a subject-oriented collection of business data identified with a particular period of time, as opposed to a transactional database which is dedicated to ongoing business activities. Data in a warehouse is normally gathered from a variety of sources (mostly various OLTP and legacy systems) and merged into a coherent whole. Data in a warehouse is usually stable, with data being added to the warehouse periodically but usually not removed. However, in some cases the oldest data may be unloaded when the newer data is added.

As opposed to data stored in OLTP systems intended to support day-to-day operations and optimized for the speed and reliability of transaction updating, data stored

in a data warehouse is intended to provide higher-level, aggregated views of the data in support of business decision making. This might include total sales by product line or region over a selectable period of time. To provide consistently fast responses to such aggregating queries, data in a data warehouse must be structured in a manner facilitating
5 its synthesis, analysis, and consolidation.

The most characteristic feature of warehoused business data is its multidimensional view of a business - so-called because the business data is organized according to major aspects and measures of the business called dimensions. The dimensions for a business may include its products, markets, profits, or time periods
10 involved. This is in contrast to a data dependencies model of the business data used to keep track of all logical relationships among all the possible data elements relevant to the business and its day-to-day operations.

A dimension may include several hierarchical levels of categories. For example, the market dimension may contain, in descending order, such categories as country,
15 region, state, and city, each category having its own number of specific instances. A hierarchical dimension reduces the total number of dimensions necessary to describe and organize the data, as compared with the situation where each category is represented by a separate dimension.

The action of viewing data in greater detail by moving down the hierarchy of categories, i.e., by moving from parent to child category, is sometimes referred to as
20 "drilling down" through the data. Quite naturally, the action of moving in the opposite direction, i.e., up the hierarchy of categories, to produce a more consolidated, higher-level view of data, is known as "drilling up" through the data. Another aspect of this 'drilling' is the ability to 'drill through' from one form of data to another where the data
25 are collected in reports having different meta-data models and/or data sources, but where the actual data are compatible (or able to be converted to a compatible form) for use in an application. Drill-through is sometimes defined as the action of navigating from one report to another report and applying the context of the source to the target. Drilling is further discussed below.

30 The vast amounts of warehoused or otherwise collected business data would be useless without software tools for its analysis. Such tools are known under the collective

name of business intelligence (BI) applications, an example of which is a suite of BI applications from Cognos Inc. (Canada). BI applications provide, among other features, data warehouse construction tools, as well as database querying, navigation and exploration tools, and also include reporting, modeling, and visualization tools. Some of these tools combine techniques of data analysis and presentation with methods for discovering hidden patterns and previously unrecognized relationships among data, an approach sometimes known as data mining.

Among BI applications previously made available to the business community, tools for visualizing business data are of particular importance. By interacting with the user graphically or pictorially, these tools have become the means of choice for consolidating and presenting vast amounts of complex information, in a manner which facilitates pattern recognition in the presented data and invites data exploration.

For example when invoking a drill-through operation from a row of the source report to a target report, a filter is constructed based on the source report row and the filter is then applied to the target report. This filter is then used by an application in requesting data from another application.

Using earlier technology it is necessary to apply some very strict rules in order to effectively drill from one application to another. The drill-through filter generated from the drill-through source is required to match up with the column names used in the target report. This column name matching is typically done through a set of intermediate files called query definitions (QD). A QD file is essentially the SQL query that can be used by a transformation program to generate a new data cube. The QD file contains the column name mapping used to generate the context filter that the target application can understand. In general, a context filter helps restrict the target report data based on the context (or some content) from the source report. For example, a context filter might specify "color='red'" because that was the value contained in a cell selected in the source report. Then the target report will only show the sales data for appliances that are red. The process is duplicated for each source/target application pair, although in each case the intermediate files may have slightly different formats, but they still served as a column name map.

For convenience, the following definitions are given for terms that are used throughout this application:

Business Intelligence Tools: Software that enables business users to see and use large amounts of complex data. The following three types of tools are referred to as

- 5 Business Intelligence Tools: Multidimensional Analysis Software - also known as Online Analytical Processing (OLAP) - Software that gives the user the opportunity to look at the data from a variety of different dimensions. Query Tools - Software that allows the user to ask questions about patterns or details in the data. Data Mining Tools - Software that automatically (and sometimes interactively) searches for significant patterns or
10 correlations in the data.

Cube (also known as Multidimensional Cube): The fundamental structure for data in a multidimensional (OLAP) system. A cube contains dimensions, hierarchies, levels, and measures. Each individual point in a cube is referred to as a cell.

- Data-Based Knowledge: Knowledge derived from data through the use of
15 Business Intelligence Tools and the process of Data Warehousing that give us a new kind of knowledge based on data. Data-based knowledge can have several advantages compared to experience/intuition-based knowledge: Improved accuracy - because it is based on so many detailed facts. More current - because the data warehousing and business intelligence tools can so quickly analyze new data. More comprehensive -
20 because so many different perspectives are available through the rapid recombination of elements from different dimensions and different levels of the data hierarchy. New insights are possible - because there are complex patterns in the data that can be discovered by data mining that would never be detected by human analysis. Less subjective - because conclusions are tied directly to the physical data.

- 25 Dimension: One of the perspectives that can be used to analyze the data in an OLAP cube. When browsing the data in a cube, it is possible to view the data from the perspective of different combinations of dimensions. For a Sales database, the dimensions might include Product_Name, Time_of_Sale, Store_Name, and Promotion_Name. Dimensions contain one or more hierarchies, which have levels for
30 drilling up and drilling down in the cube. When a dimension has just one hierarchy, which is quite common, people often refer to the dimension itself as having levels.

Hierarchy: Organization of data into a logical tree structure. Dimensions can have one or more hierarchies. A Time dimension, for example, might have a Calendar hierarchy and a Fiscal hierarchy. Hierarchies contain levels, which organize data in a logical structure. It is the combination of multidimensional data with a hierarchical view
5 given by Business Intelligence Software that allows users to grapple successfully with large amounts of data. If each member in a level has 5 to 10 children that are members at the next lower level, the user has a better chance of understanding the significance of the data. Moving between the levels of a hierarchy is called drilling up and drilling down.

Level: The hierarchies in dimensions have levels which can be used to view data
10 at various levels of detail. A Time dimension might have levels for Year, Quarter, Month, and Day. A Product dimension might have levels for Product Family, Product Category, Product Sub-category, and Product Name. A Customer Geography dimension might have levels for Region, Country, District, State, City, and Neighborhood.

Member: One of the data points for a level of a hierarchy of a dimension. Some of
15 the members of the Month level of the Time dimension are January, February, March, and April.

Measure: A numeric value stored in a fact table and in an OLAP cube. Sales Count, Sales Price, Cost, Discount, and Profit could all be measures in an OLAP cube.

Multidimensional Analysis - also known as On-Line Analytical Processing
20 (OLAP): A process of analysis that involves organizing and summarizing data in a multiple number of dimensions. People can comprehend a far greater amount of information if that information is organized into dimensions and into hierarchies. The wide use of spreadsheets and graphs illustrates the need for people to have their information organized. A spreadsheet is a two-dimensional analysis tool. If a person
25 could comprehend ten individual facts, they might possibly comprehend 100 facts if they were arranged in a spreadsheet. Simplistically, if three or four or five dimensions were displayed, the amount of information that could be comprehended might increase exponentially - to 1000 facts, 10,000 facts, and 100,000 facts respectively.

Multidimensional data is also organized hierarchically, allowing users to "drill down" for
30 more detailed information, "drill up" to see a broader, more summarized view, and "slice and dice" to dynamically change the combinations of dimensions that are being viewed.

Summary of the Invention

The present invention is designed to overcome at least some of the problems found in earlier systems, and also provide other advantages as will be clear from the following description.

5 In the invention, a drill-through path is defined by a collection of parameter mappings that map context elements from the drill-through source (report) to the inputs of the target. In addition to defining the target/source mapping, these parameter mappings also define any translation, filtering, or formatting that may be required.

A parameter is a way of providing 'public' access to a column for external
10 applications by giving it a name. More generally parameters define an input or output to a drill-through source or target. Additional parameters may be used to select and define conversion functions for the parameter mappings in the drill-through process so that the invention also acts as a means for defining a data converter or filter.

According to one aspect of the present invention, there is provided a database
15 interface for generating drill-throughs of various data cubes, derived from different report generating applications. The interface comprises: (a) means for accepting a request from a user for data; (b) means for translating the request into a drill-through selected from a plurality of possible drill-through paths between a source and a target; (c) means for applying one or more parameters to the selected drill-through path to produce a valid
20 drill-through path and to transfer the requested data over the valid drill-through path to an application; and (d) display means for displaying the requested data to the user.

According to another aspect of the present invention, there is provided a computer-based method for obtaining data from one or more compatible data sources for use within applications implementing a decision support system, the method comprising the steps of,
25 in a business modeling tool before using a business intelligence application, (a) modeling a mapping of data between a source and a target to produce one or more possible drill-through paths between the source and the target, each drill-through path having one or more parameter mappings, in a business intelligence application, using a report authoring tool, (b) accepting a request from a user for data, (c) translating the request into a drill-
30 through path selected from the possible drill-through paths between the source and the target, (d) applying one or more parameters to the selected drill-through path to produce a

valid parameter mapping and transferring the requested data over the valid parameter mapping to an application, and (e) displaying the requested data to a user.

According to another aspect of the present invention, there is provided a computer-based system for obtaining data from one or more compatible data sources for use within applications implementing a decision support system, the system comprising, in a business modeling tool, before using a business intelligence application, (a) means for modeling a mapping of data between a source and a target to produce one or more possible drill-through paths between the source and the target, each drill-through path having one or more parameter mappings, and in a business intelligence application, using a report authoring tool, (b) means for accepting a request from a user for data, (c) means for translating the request into a drill-through path selected from the possible drill-through paths between the source and the target, (d) means for applying one or more parameters to the selected drill-through path to produce a valid parameter mapping and to transfer the requested data over the valid parameter mapping to the application, and (e) display means for displaying the requested data to a user.

According to another aspect of the present invention, there is provided an apparatus for obtaining data from one or more compatible data sources for use within applications implementing a decision support system, the apparatus comprising, in a business modeling tool, (a) means for modeling a mapping of data between a source and a target to produce one or more possible drill-through paths between the source and the target, each drill-through path having one or more parameter mappings, and in a business intelligence application, using a report authoring tool, (b) means for accepting a request from a user for data, (c) means for translating the request into a drill-through path selected from the one or more possible drill-through paths between the source and the target, (d) means for applying one or more parameters to the selected drill-through path to produce a valid parameter mapping and to transfer the requested data over the valid parameter mapping to the application, and (e) display means for displaying the requested data to the user.

A further understanding of other features, aspects and advantages of the present invention will be realized by reference to the following description, appended claims and accompanying drawings.

Brief Description of the Drawings

A preferred embodiment(s) of the invention will now be described with reference to the following figures:

Figure 1 shows the computer and related systems and connections within which
5 the invention may be practiced;

Figure 2 shows a simple situation in which a single parameter mapping defines a single item;

Figure 3 shows a more complex example, in which there are several items, each defined by one or more parameter mappings;

10 Figure 4 illustrates that there may be different functions related to each bi-directional parameter mapping in a drill-through path;

Figure 5 shows the necessary parameter mappings and drill-through paths for a situation where two business models must be integrated;

Figure 6 is a more complex example, relating three business models; and

15 Figure 7 is a further example, also showing the relationship with a Dimension Map.

Detailed Description of the Preferred Embodiment(s)

Typically, embodiments of the invention are practiced within a network of computers and related equipment as illustrated in Figure 1. Data is stored in a data
20 warehouse comprising files servers 100, 102, 104 and storage devices 106, 108, with their related software applications (not shown), accessed from workstations 110, 111, 112 over a network 120, the workstations also including related software applications (not shown).

Turning now to the Figure 2, a first preferred embodiment is shown in which an
25 item 210, representing a report column, or parameter, or some other expression from a drill-through source or target, has a parameter 200 exposed. The parameter mapping 240 from the item includes two optional mapping functions 220, 230 which may translate the data from the outside world (using the parameter 200) to the item and vice-versa.

Several optional functions made feasible by other preferred embodiments are
30 describe below. One function already mentioned is the ability to translate or convert the

data during the drill-through operation, and to select the actual conversion by means of a parameter. Earlier applications not incorporating embodiments of the present invention were only able to perform data conversion or manipulation of drill-through parameters within the report-producing application, and the required functions were stored with the report schema. This meant that the report was often used specifically for one drill-through since the report column used in the context filter would only make sense for that particular drill-through. By enhancing the drill-through service so the input and output conversion functions are part of the drill-through model it is possible to decouple them from the report itself (or its description and schema). In other words, any data conversions required for a drill-through path are encapsulated in the drill-through service, rather than in the report (and its schema). This is illustrated in the Figure 3, in which a drill-through source or target is generalized to a collection of parameters 300, 320, 340, 360 as shown each related (or mapped) over a parameter mapping including a function 305, 325, 345, 365 respectively to one of the items 310, 330, 350. Note that more than one parameter can map to a single drill-through item as is shown in the figure 3, where parameter 340 and parameter 360 both map to item 350 through their respective functions in parameter mappings 345, 365.

Another feature made possible by embodiments of the present invention is a drill-through path. A drill-through path is a collection of source and target parameter mappings. The drill-through path effectively couples the source and target objects together. A single drill-through target can have many drill-through sources and vice-versa. This aspect is illustrated in Figure 3, described earlier.

Figure 4 shows an embodiment in which several bi-directional drill-through mappings are illustrated, 400-402, 420-422, 430-432, 440-442, each having different mapping functions for each direction, 404 and 406, 424 and 426, 444 and 446, and 464 and 466.

All of the various functions and mappings are captured and modeled using a single drill-through administration application. One preferred embodiment provides a single drill-through service that may be queried for drill-through information such as a list of targets from a given source or drill-through actions to be performed.

The various concepts that are used in the new drill-through mechanism of the present invention will be described below.

The invention solves several other problems found in earlier implementations. Since the drill-through source and targets are generalized as a collection of input and output parameters, it is no longer important to know in advance the actual drill-through target. This means, for example, that cubes built from data extracted from database "A" can still drill-through to a report whose data comes from database "B". It is the modeler's responsibility to ensure that the drill-through paths from the source and targets are properly matched up and the appropriate conversion functions are made available in the drill-through service.

Provision of the new drill-through mechanism of the present invention allows the drill-through model to be used in defining an Internet Uniform Resource Locator (URL) or a Hypertext Markup Language (HTML) FORM template, thereby providing for standard procedures in accessing data over the Internet. Therefore, in some preferred embodiments, a drill-through to any web-based target makes use of the generic ability to add parameters to a URL or FORM template. For example, in a case where a template is a string containing one or more predefined drill-through parameter placeholders, defined as <P1>, <P2>... etc., a drill-through target is defined as a URL, using a well-known construction, for example:

`http://myserver/path/app.cgi?params=(<P1>,<P2>,<P3>)`

Consider the case of a report that contains 3 columns named C1, C2, and C3. A drill-through path is defined that maps this source with the URL target. In this case the drill-through path is (C1->P1, C2->P2, C3->P3). Drilling from the source report at a specific row to the target means that the row values for C1, C2 and C3 will be substituted into the URL at the P1, P2, and P3 placeholders, thereby producing a valid URL, for example:

`http://myserver/path/app.cgi?params=(C1, C2, C3)`

A meta-data model is the model on which the reports are based. It is sometimes convenient to work with the underlying metadata, rather than the reports themselves. Therefore the concept of defining one or more drill-through paths between meta-data models will be described below. In the embodiments described above a drill-through

model designer is able to define a drill-through path from one report to another. The designer (user) may optionally define any parameters to be used in the context filter when navigation from one report to another. In addition, drill-through services embodying the invention support the definition of one or more drill-through paths between the different meta-data models. This has the advantage that reports based on the meta-data models can make implicit use of the already defined drill-through paths. The invention therefore affords a generalized way of assigning drill-through paths between reports without having to explicitly define drill-through paths between all reports. This is done by providing a list of potential drill-through paths between the query items of a model.

Various aspects of the use of the present invention are more fully explained by reference to the following examples.

Considering a situation where there are two different meta-data models. The first meta-data model is based on the "Sales" database and is used to allow the report writer to generate reports on all sales related information. The other meta-data model is based on the "Human Resources" database and allows report writers to generate reports of Personnel information. In some circumstances it is desirable to allow managers to drill-through from sales-based reports to personnel-based reports in order to answer questions such; "Is this individual's salary commensurate with his sales record?"

We next describe in more detail how these kinds of queries are supported using embodiments of the present invention. In the Figure 5, the two database models, Sales 500 and Human Resources 520 are shown, together with relationships (or drill-through paths) 540 represented by lines drawn between the related items representing their parameter mappings P1 - P5. P1 - P5 are determined and suitably defined by a trained modeler experienced in the use of such systems. The Sales model includes two separate tables, namely SalesEmployee 502 and SalesMetrics 510, and the Human Resources Model includes a PersonalInfo table 521, and a SalaryInfo table 530. The relationships within 540 represent the parameter mapping from one model to the other as previous described with reference to Figures 3 and 4. In this case, the following parameter mappings are defined:

P1 between [SalesEmployee].[EmployeeNum] 503 and
[PersonalInfo].[EmpNum] 522.

P2 between [SalesEmployee].[Manager] 505 and [PersonalInfo].[Name] 523.

P3 between [SalesEmployee].[EmployeeNum] 503 and
[SalaryInfo].[EmpNum] 531.

P4 between [SalesEmployee].[Name] 504 and [PersonalInfo].[Name] 523.

5 P5 between [SalesMetrics].[EmployeeNum] 511 and
[PersonalInfo].[EmpNum] 522.

Assume a source report is based on the Sales model (Sales) with the following columns;

10 [SalesEmployee].[Manager], [SalesEmployee].[Name],
total([SalesMetrics].[Revenue])

Further assume that the target report is to be based on the Human Resources model (HR info) with the following report columns;

[PersonalInfo].[Name], [PersonalInfo].[Title], [PersonalInfo].[Salary],
[PersonalInfo].[LastRaiseDate]

15 Note that it is not necessary to have parameters assigned to all of the query items. In this example [PersonalInfo].[Title] 524, and [SalaryInfo].[LastRaiseDate] 532 are not assigned parameters.

For the sake of simplicity no parameter mapping functions are illustrated in this example. Their addition and effect when required will be obvious to one skilled in the art.

20 The report authoring tool first asks the drill-through service for a list of potential parameter mappings in drill-through path that could be used to drill-through from the source report (Sales) 500 to the target report (HR Info) 520. The drill-through service employs the following algorithm to determine the mappings:

1. Create a list of parameters (query items) from the source report 500 and target
25 report 520.
2. For each source parameter determine a parameter mapping that maps the parameter to the target. All defined parameter mappings are collected as a single drill-through path
3. If more than one parameter mapping points to the same target parameter then
30 duplicate the parameter mapping one for each duplicate target mapping. For example, in Figure 5 there are potentially two parameter mappings pointing to the

same target [PersonalInfo].[Name] 523. As a result, two separate parameter mappings are created from the sources [SalesEmployee].[Name] 504 and [SalesEmployee].[Manager] 505. These parameter mappings are P2 and P4 within the drill-thorough path 540. This duplication prevents any conflicts when the filter is formed.

4. Continue duplicating until all the parameter mappings for the drill-through path point to unique target parameters.

In this example, after completion of the above algorithm, the following parameter mappings from among those possible 540 will be returned from the drill-through service:

- Parameter mapping #1 (P4) and
- Parameter mapping #2 (P2).

Using these parameter mappings it is possible to drill-through from the sales report to the HR report and filter on either the salesperson's name or on the manager's name. The parameterised drill-through service of the present invention therefore allows the user to obtain the answer to either or both of the two questions:

1. "How much am I paying a sales person relative to their generated revenue?" and
2. "How much am I paying a manager relative to the total revenue of the sales people being managed?"

Only a single parameter mapping for each drill-through path is used here since only one parameter is matched from source to target. Generally, a drill-through path may contain any number of applicable parameter mappings.

Since it cannot be known in advance what parts of the report a model may use, a drill-through path is fully defined during the modeling phase. Therefore a collection of parameter mappings (P1- P5) 540 is provided from which the actual drill-through path can be built dynamically depending on what columns (parameters) are available in the source and target reports.

It is possible to have more than one parameter mapping generated, as shown here, so that a drill-through path selection must be made by the user.

Expanding on this general model of drill-through usage, it is also possible to define general drill-through paths from a meta-data model to specific reports. Figure 6

shows a Sales model 600 having two drill-through paths 630, 635 defined to two separate reports: an Order Detail Report 620 and an Employee Expenses Report 625.

One of the parameter mappings of the "OrderDetail" drill-through path 630 maps [SalesMetrics].[OrderNum] 605 to report column "Order Number" 621. The "Expenses" drill-through path 635 contains a single parameter mapping on [SalesEmployee].[EmployeeNum] 602 to report column "Employee Number" 626. In some cases other parameters are available to be used in other instances of using the drill-through service. In some cases it is up to the drill-through target report to dispense with unnecessary filtering as part of any optimisation.

Once these drill-through services indicated by the paths 630, 635 have been defined, then any report that included either or both of the Sales Model 600 items [SalesModel].[SalesEmployee].[Name] 603 and [SalesModel].[SalesMetrics].[OrderNumber] 605 can drill-through to the OrderDetail report 620. Similarly any report that includes a report column based on [SalesEmployee].[EmployeeNum] 602 can drill-through to the Employee Expenses report 625.

This next example describes the drill-through parameter mapping to and from OLAP sources to models for relational reporting tools.

Referring to the Figure 7, a cube 710 (represented in this diagram as being derived from a dimension map 700 by a suitable transformation tool 705) has a set of parameters which are the dimension levels. Each level is represented by (or exposed as) a parameter.

In one case a drill-through path 790 is defined from the cube 710 to the report OrderDetail Report 760, the path having two parameters mappings (OrderNum) 761, (Outlet) 762.

In another case the cube parameters 710 are mapped to a model. This is a generalized case which is similar to the model to model case previously described. Drill through paths 795 and their related parameter mappings 781 are defined between the cube 710 and the Model 780. As a consequence, any reports that are based on the model can drill through into the cube assuming that the required parameters (columns) are included in the report.

It has been found that significant savings are realised in terms of the number and the updating of drill-through paths, when compared to the situation before the availability of parameterised drill-through paths and their associated parameter mappings.

While various embodiments of the present invention have been described above,
5 it should be understood that they have been presented by way of example only, and not limitation. It will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the spirit and scope of the invention as defined in the appended claims. Thus, the breadth and scope of the present invention should not be limited by any of the above-described exemplary embodiments,
10 but should be defined only in accordance with the attached or associated claims and their equivalents.

The invention can be implemented in digital electronic circuitry or in computer hardware, firmware, and software or in combinations thereof. Apparatus of the present invention can be implemented in a computer program product tangibly embodied in a
15 machine-readable storage device for execution by a programmable processor. The method of the present invention can be performed by a programmable processor executing a program of instructions to perform the functions of the invention by operating on input data and generating output.